

Claims:

1 1. A system for automatically processing a textual document, comprising:
2 means for identifying occurrences of a piece of text having a predetermined format
3 in the textual document;
4 means for determining a depth-of-treatment value for each piece of formatted text in
5 the textual document, the depth-of-treatment value indicating a depth of treatment in the
6 textual document afforded to the particular piece of formatted text;
7 means for associating an abstract with each piece of formatted text in the textual
8 document; and
9 means for generating a data record containing each identified piece of formatted text
10 from the textual document, the depth-of-treatment value and the abstract associated with
11 each piece of formatted text in the textual document.

1 2. The system of Claim 1, wherein the textual document comprises the written
2 decision of a court for a particular legal case and wherein each piece of formatted text in the
3 textual document comprises a citation to another legal authority, the citation indicating the
4 location of the other legal authority.

1 3. The system of Claim 2, wherein each piece of formatted text further
2 comprises a quotation from said other legal authority.

1 4. The system of Claim 3 further comprises means for identifying a quotation
2 which is a candidate for being associated with a citation in the textual document, and means

3 for verifying a source document for each candidate quotation in the textual document to
4 generate verified quotation information.

1 5. The system of Claim 4, wherein the verifying means further comprises means
2 for separating a quotation into one or more segment based on the location of a predetermined
3 symbol in the quotation, means for calculating a word frequency value for each word in each
4 segment, means for selecting a predetermined number of words in each segment, each
5 selected word having a high word frequency value, means for generating a template of the
6 quotation indicating the position of each of the selected words in each segment, and means
7 for comparing the generated template to the text of the quotation in the source document.

1 6. The system of Claim 2 further comprising means for determining a phrase in
2 the textual document indicating a negative treatment of another legal authority comprising
3 means for parsing the textual document to identify occurrences of a predetermined word
4 stem in the textual document indicating that the reasoning of a prior written decision of a
5 court is not proper and means for verifying said identified word stems to generate a list of
6 overrulings.

1 7. The system of Claim 6, wherein said parsing means comprises means for
2 identifying each occurrence of a particular tense of the word stem, means for gathering
3 words adjacent to each occurrence of the particular tense of the word stem, and means for
4 selecting a negative treatment phrase based on the gathered words.

1 8. The system of Claim 2 further comprising means for determining a phrase in
2 the textual document indicating a negative treatment of another legal authority comprising
3 means for parsing the textual document to identify occurrences of a predetermined word in
4 the textual document indicating that the reasoning of a prior written decision of a court is not
5 proper and means for verifying said identified word stems to generate a list of negative
6 treatments.

1 9. The system of Claim 2, wherein said formatted text identification means
2 further comprises means for identifying numerical digits in the textual document, means for
3 locating, in proximity to the numerical digits, each occurrence of an abbreviation in the
4 textual document, the abbreviation indicating the location of a written decision referred to by
5 the piece of formatted text, and means for gathering the numerical digits and the
6 abbreviations to identify a citation.

1 10. The system of Claim 1, wherein said depth-of-treatment generating means
2 comprises means for generating a depth symbol for each piece of formatted text in the
3 textual document indicating a potential significance of the piece of formatted text.

1 11. The system of Claim 10, wherein said depth symbol generating means
2 comprises means for selecting a type of each occurrence of a piece of formatted text from a
3 predetermined set of types, means for determining, for each piece of formatted text, the
4 number of each type of occurrence of the piece of formatted text, and means for assigning
5 the depth symbol to each piece of formatted text in the textual document based on the total

6 number of occurrences of the text, the number of each type of occurrence of the text, and any
7 verified quotations associated with the piece of formatted text.

1 12. The system of Claim 11, wherein said depth symbol comprises a symbol
2 having a predetermined number of elements corresponding to the depth-of-treatment value.

1 13. The system of Claim 4, wherein the subject matter classification means
2 comprises means for identifying sentences adjacent to the occurrences of each piece of
3 formatted text, means for executing a query for each piece of formatted text based on the
4 identified sentences and a verified quotation associated with the piece of formatted text to
5 generate one or more subject matter classification candidates and a belief score for each
6 subject matter classification candidate, and means for selecting one or more of the subject
7 matter classification candidates having a belief score above a predetermined threshold to
8 associate with the piece of formatted text.

1 14. The system of Claim 3, wherein said formatted text identification means
2 comprises means for parsing the textual document into one or more paragraphs, means for
3 identifying, for each paragraph, one or more tokens which indicate the presence of a
4 quotation in the sentence, and means for determining, based on the identified tokens,
5 quotations within the textual document.

1 15. A method for automatically processing a textual document, comprising:

identifying occurrences of a piece of text having a predetermined format in the textual document;

determining a depth-of-treatment value for each piece of formatted text in the textual document, the depth-of-treatment value indicating a depth of treatment in the textual document afforded to the particular piece of formatted text;

associating an abstract with each piece of formatted text in the textual document; and

generating a data record containing each identified piece of formatted text from the textual document, the depth-of-treatment value and the abstract associated with each piece of formatted text in the textual document.

16. The method of Claim 15, wherein the textual document comprises the written decision of a court for a particular legal case and wherein each piece of formatted text in the textual document comprises a citation to another legal authority, the citation indicating the location of the other legal authority.

17. The method of Claim 16, wherein each piece of formatted text further comprises a quotation from said other legal authority.

18. The method of Claim 17 further comprises identifying a quotation which is a candidate for being associated with a citation in the textual document, and verifying a source document for each candidate quotation in the textual document to generate verified quotation information.

1 19. The method of Claim 18, wherein the verifying further comprises separating a
2 quotation into one or more segment based on the location of a predetermined symbol in the
3 quotation, calculating a word frequency value for each word in each segment, selecting a
4 predetermined number of words in each segment, each selected word having a high word
5 frequency value, generating a template of the quotation indicating the position of each of the
6 selected words in each segment, and comparing the generated template to the text of the
7 quotation in the source document.

1 20. The method of Claim 16 further comprising determining a phrase in the
2 textual document indicating a negative treatment of another legal authority comprising
3 parsing the textual document to identify occurrences of a predetermined word stem in the
4 textual document indicating that the reasoning of a prior written decision of a court is not
5 proper and verifying said identified word stems to generate a list of overrulings.

1 21. The method of Claim 20, wherein said parsing comprises identifying each
2 occurrence of a particular tense of the word stem, gathering words adjacent to each
3 occurrence of the particular tense of the word stem, and selecting a negative treatment phrase
4 based on the gathered words.

1 22. The method of Claim 16 further comprising determining a phrase in the
2 textual document indicating a negative treatment of another legal authority comprising
3 parsing the textual document to identify occurrences of a predetermined word in the textual

4 document indicating that the reasoning of a prior written decision of a court is not proper and
5 verifying said identified word stems to generate a list of negative treatments.

1 23. The method of Claim 16, wherein said formatted text identification further
2 comprises identifying numerical digits in the textual document, locating, in proximity to the
3 numerical digits, each occurrence of an abbreviation in the textual document, the
4 abbreviation indicating the location of a written decision referred to by the piece of
5 formatted text, and gathering the numerical digits and the abbreviations to identify a citation.

6

1 24. The method of Claim 15, wherein said depth-of-treatment generating
2 comprises generating a depth symbol for each piece of formatted text in the textual
3 document indicating a potential significance of the piece of formatted text.

1 25. The method of Claim 24, wherein said depth symbol generating comprises
2 selecting a type of each occurrence of a piece of formatted text from a predetermined set of
3 types, determining, for each piece of formatted text, the number of each type of occurrence
4 of the piece of formatted text, and assigning the depth symbol to each piece of formatted text
5 in the textual document based on the total number of occurrences of the text, the number of
6 each type of occurrence of the text, and any verified quotations associated with the piece of
7 formatted text.

1 26. The method of Claim 25, wherein said depth symbol comprises a symbol
2 having a predetermined number of elements corresponding to the depth-of-treatment value.
3

1 27. The method of Claim 18, wherein the subject matter classification comprises
2 identifying sentences adjacent to the occurrences of each piece of formatted text, executing
3 a query for each piece of formatted text based on the identified sentences and a verified
4 quotation associated with the piece of formatted text to generate one or more subject matter
5 classification candidates and a belief score for each subject matter classification candidate,
6 and selecting one or more of the subject matter classification candidates having a belief
7 score above a predetermined threshold to associate with the piece of formatted text.

1 28. The method of Claim 17, wherein said formatted text identification
2 comprises parsing the textual document into one or more paragraphs, identifying, for each
3 paragraph, one or more symbols which indicate the presence of a quotation in the sentence,
4 and determining, based on the identified symbols, quotations within the textual document.
5

1 29. A method for automatically processing a textual document, comprising:
2 parsing the text of the textual document into one or more paragraphs;
3 identifying, within each paragraph, one or more symbols which indicates the
4 presence of a quotation text in the paragraph;

5 determining, based on the identified symbols, a candidate quotation within the
6 textual document and a potential source document of the quotation; and
7 matching the candidate quotation with quotation text in the potential source
8 document to verify the source of the quotation.

1 30. The method of Claim 29, wherein the matching further comprises separating
2 a quotation into one or more segment based on the location of a predetermined symbol in the
3 quotation, calculating a word frequency value for each word in each segment, selecting a
4 predetermined number of words in each segment, each selected word having a high word
5 frequency value, generating a template of the quotation indicating the position of each of the
6 selected words in each segment, and comparing the generated template to the text of the
7 quotation in the source textual document.

1 31. A system for automatically processing a textual document, comprising:
2 means for parsing the text of the textual document into one or more paragraphs;
3 means for identifying, within each paragraph, one or more symbols which indicates
4 the presence of a quotation text in the paragraph;
5 means for determining, based on the identified symbols, a candidate quotation within
6 the textual document and a potential source document of the quotation; and
7 means for matching the candidate quotation with quotation text in the potential
8 source document to verify the source of the quotation.

1 32. The system of Claim 31, wherein the matching means further comprises
2 means for separating a quotation into one or more segment based on the location of a
3 predetermined symbol in the quotation, means for calculating a word frequency value for
4 each word in each segment, means for selecting a predetermined number of words in each
5 segment, each selected word having a high word frequency value, means for generating a
6 template of the quotation indicating the position of each of the selected words in each
7 segment, and means for comparing the generated template to the text of the quotation in the
8 source textual document.

1 33. An automatic text processing system, comprising:
2 means for identifying each occurrence of a piece of text having a predetermined
3 format in the textual document; and
4 means for determining a value for each piece of formatted text in the textual
5 document indicating a depth-of-treatment of the piece of formatted text in the textual
6 document, the determining means comprising means for selecting a type of each occurrence
7 of a piece of formatted text from a predetermined set of types, means for determining, for
8 each piece of formatted text, the total number of occurrences of the piece of formatted text in
9 the textual document and the number of each type of occurrence for each piece of formatted
10 text, and means for assigning a depth-of-treatment value to each piece of formatted text in
11 the textual document based on the total number of occurrences of the formatted text and the
12 number of each type of occurrence of the formatted text.

1 34. The system of Claim 33 further comprising means for locating a quotation in
2 the textual document and means for verifying a source document for a quotation in the
3 document to generate a verified quotation, and wherein said assigning means further
4 comprises means for determining a depth-of-treatment symbol based on any verified
5 quotations associated with each piece of formatted text.

1 35. The system of Claim 34, wherein the verifying means further comprises
2 means for separating a quotation into one or more segment based on the location of a
3 predetermined symbol in the quotation, means for calculating a word frequency value for
4 each word in each segment, means for selecting a predetermined number of words in each
5 segment, each selected word having a high word frequency value, means for generating a
6 template of the quotation indicating the position of each of the selected words in each
7 segment, and means for comparing the generated template to the text of the quotation in the
8 source document.

1 36. The system of Claim 33, wherein the depth-of-treatment value generating
2 means comprises means for generating a depth symbol that is associated with each piece of
3 formatted text in the textual document.

1 37. The system of Claim 36, wherein said depth symbol comprises a symbol
2 having a predetermined number of elements corresponding to the depth-of-treatment value.

1 38. An automatic text processing method, comprising:
2 identifying each occurrence of a piece of text having a predetermined format in the
3 textual document; and
4 determining a value for each piece of formatted text in the textual document
5 indicating a depth-of-treatment of the piece of formatted text in the textual document, the
6 determining comprising selecting a type of each occurrence of a piece of formatted text from
7 a predetermined set of types, determining, for each piece of formatted text, the total number
8 of occurrences of the piece of formatted text in the textual document and the number of each
9 type of occurrence for each piece of formatted text, and assigning a depth-of-treatment value
10 to each piece of formatted text in the textual document based on the total number of
11 occurrences of the formatted text and the number of each type of occurrence of the formatted
12 text. .

1 39. The method of Claim 38 further comprising locating a quotation in the textual
2 document and verifying a source document for a quotation in the document to generate a
3 verified quotation, and wherein said assigning further comprises determining a depth-of-
4 treatment symbol based on any verified quotations associated with each piece of formatted
5 text. .

1 40. The method of Claim 39, wherein the verifying further comprises separating a
2 quotation into one or more segment based on the location of a predetermined symbol in the
3 quotation, calculating a word frequency value for each word in each segment, selecting a
4 predetermined number of words in each segment, each selected word having a high word

frequency value, generating a template of the quotation indicating the position of each of the selected words in each segment, and comparing the generated template to the text of the quotation in the source document. .

41. The method of Claim 38, wherein the depth-of-treatment value generating comprises generating a depth symbol that is associated with each piece of formatted text in the textual document. .

42. The method of Claim 41, wherein said depth symbol comprises a symbol having a predetermined number of elements corresponding to the depth-of-treatment value. .

43. A system for automatically processing a textual document, comprising:
means for identifying each occurrence of a piece of text having a predetermined format in the textual document;
means for locating and verifying a quotation in the document to generate a verified quotation; and
means for associating a subject matter classification with each piece of formatted text in the textual document, comprising means for identifying sentences adjacent to the occurrences of each piece of formatted text, means for executing a query for each piece of formatted text based on the identified sentences and a verified quotation associated with the piece of formatted text to generate one or more subject matter classification candidates and a belief score for each subject matter classification candidate, and means for selecting one or

12 more subject matter classification candidates having a belief score above a predetermined
13 threshold to associate with the piece of formatted text..

1 44. The system of Claim 43 further comprising means for determining a depth-
2 of-treatment value for each piece of formatted text in the text of the textual document
3 indicating the level of treatment afforded to the piece of formatted text by the textual
4 document. .

1 45. The system of Claim 44, wherein said depth-of-treatment value determining
2 comprises means for selecting a type of each occurrence of a piece of formatted text from a
3 predetermined set of types, means for determining, for each piece of formatted text, the total
4 number of occurrences of the piece of formatted text and the number of each type of
5 occurrence for each piece of formatted text, and means assigning the depth-of-treatment
6 value to each piece of formatted text in the textual document based on the total number of
7 occurrences of the formatted text, the number of each type of occurrence of the formatted
8 text. .

1 46. The system of Claim 43 further comprising means for locating a quotation in
2 the textual document and means for verifying a quotation in the document to generate a
3 verified quotation, and wherein said assigning means further comprises means for
4 determining a depth symbol based on any verified quotations associated with the piece of
5 formatted text.

1 47. The system of Claim 46, wherein the verifying means further comprises
2 means for separating a quotation into one or more segment based on the location of a
3 predetermined symbol in the quotation, means for calculating a word frequency value for
4 each word in each segment, means for selecting a predetermined number of words in each
5 segment, each selected word having a high word frequency value, means for generating a
6 template of the quotation indicating the position of each of the selected words in each
7 segment, and means for comparing the generated template to the text of the quotation in the
8 source textual document. .

1 48. The system of Claim 44, wherein the depth-of-treatment value generating
2 means comprises means for generating a depth symbol that is associated with each piece of
3 formatted text in the textual document.

1 49. The system of Claim 48, wherein said depth symbol generating means
2 comprises means for selecting a type of each occurrence of a piece of formatted text from a
3 predetermined set of types, means for determining, for each piece of formatted text, the
4 number of each type of occurrence of the formatted text, and means for assigning the depth
5 symbol to each piece of formatted text in the textual document based on the total number of
6 occurrences of the formatted text, the number of each type of occurrence of the formatted
7 text, and any verified quotations associated with the piece of formatted text. .

1 50. The system of Claim 49, wherein said depth symbol comprises a symbol
2 having a predetermined number of element corresponding to the depth-of-treatment value. .
3

1 51. A method for automatically processing a textual document, comprising:
2 identifying each occurrence of a piece of text having a predetermined format in the
3 textual document;
4 locating and verifying a quotation in the document to generate a verified quotation;
5 and
6 associating a subject matter classification with each piece of formatted text in the
7 textual document, comprising means for identifying sentences adjacent to the occurrences of
8 each piece of formatted text, means for executing a query for each piece of formatted text
9 based on the identified sentences and a verified quotation associated with the piece of
10 formatted text to generate one or more subject matter classification candidates and a belief
11 score for each subject matter classification candidate, and means for selecting one or more
12 subject matter classification candidates having a belief score above a predetermined
13 threshold to associate with the piece of formatted text..

1 52. The method of Claim 51 further comprising determining a depth-of-treatment
2 value for each piece of formatted text in the text of the textual document indicating the level
3 of treatment afforded to the piece of formatted text by the textual document. .

1 53. The method of Claim 52, wherein said depth-of-treatment value determining
2 comprises selecting a type of each occurrence of a piece of formatted text from a
3 predetermined set of types, determining, for each piece of formatted text, the total number of
4 occurrences of the piece of formatted text and the number of each type of occurrence for
5 each piece of formatted text, and assigning the depth-of-treatment value to each piece of
6 formatted text in the textual document based on the total number of occurrences of the
7 formatted text, the number of each type of occurrence of the formatted text. .

1 54. The method of Claim 51 further comprising locating a quotation in the textual
2 document and verifying a quotation in the document to generate a verified quotation, and
3 wherein said assigning further comprises determining a depth symbol based on any verified
4 quotations associated with the piece of formatted text. .

1 55. The method of Claim 54, wherein the verifying further comprises separating a
2 quotation into one or more segment based on the location of a predetermined symbol in the
3 quotation, calculating a word frequency value for each word in each segment, selecting a
4 predetermined number of words in each segment, each selected word having a high word
5 frequency value, generating a template of the quotation indicating the position of each of the
6 selected words in each segment, and comparing the generated template to the text of the
7 quotation in the source textual document. .

1 56. The method of Claim 52, wherein the depth-of-treatment value generating
2 comprises generating a depth symbol that is associated with each piece of formatted text in
3 the textual document. .

1 57. The method of Claim 56, wherein said depth symbol generating comprises
2 selecting a type of each occurrence of a piece of formatted text from a predetermined set of
3 types, means for determining, for each piece of formatted text, the number of each type of
4 occurrence of the formatted text, and assigning the depth symbol to each piece of formatted
5 text in the textual document based on the total number of occurrences of the formatted text,
6 the number of each type of occurrence of the formatted text, and any verified quotations
7 associated with the piece of formatted text. .

1 58. The method of Claim 57, wherein said depth symbol comprises a symbol
2 having a predetermined number of element corresponding to the depth-of-treatment value.
3

1 59. A system for automatically processing a written opinion of a legal case,
2 comprising:
3 means for identifying each occurrence of a piece of text having a predetermined
4 format, the formatted text being a citation in the written opinion;
5 means for determining a value for each citation indicating a depth-of-treatment
6 afforded to the particular citation in the textual document;
7 means for associating a subject matter classification with each citation; and

8 means for generating a data record containing each citation, the depth-of-treatment
9 value for each citation, and the subject matter classification associated with each citation.

1 60. The system of Claim 59, wherein each piece of formatted text further
2 comprises a quotation from another prior written decision of a court. .

1 61. The system of Claim 60 further comprises means for verifying a source for
2 each quotation in the textual document to generate verified quotation information. .

1 62. The system of Claim 61, wherein the verifying means further comprises
2 means for separating a quotation into one or more segment based on the location of a
3 predetermined symbol in the quotation, means for calculating a word frequency value for
4 each word in each segment, means for selecting a predetermined number of words in each
5 segment, each selected word having a high word frequency value, means for generating a
6 template of the quotation indicating the position of each of the selected words in each
7 segment, and means for comparing the generated template to the text of the quotation in the
8 source textual document. .

1 63. The system of Claim 59 further comprising means for determining the
2 negative history associated with the citations in the written opinion, the negative treatment
3 determining means comprising means for parsing the textual document to identify
4 occurrences of a predetermined word stem in the textual document indicating that the

5 reasoning of a prior written decision of a court is not proper and means for verifying said
6 identified word stems to generate a list of overrulings. .

1 64. The system of Claim 63, wherein said parsing means comprises means for
2 identifying each occurrence of a particular tense of the word stem, means for gathering
3 words adjacent to each occurrence of the particular tense of the word stem, and means for
4 selecting a negative treatment phrase based on the gathered words..

1 65. The system of Claim 59 further comprising means for determining a negative
2 treatment phrase in the textual document comprising means for parsing the textual document
3 to identify occurrences of a predetermined word in the textual document indicating that the
4 reasoning of a prior written decision of a court is not proper and means for verifying said
5 identified word stems to generate a list of negative treatments. .

1 66. The system of Claim 60, wherein said citation identification means comprises
2 means for locating each occurrence of a reporter name in the textual document, and means
3 for gathering words adjacent to the reporter name to identify a beginning and an end of the
4 citation..

1 67. The system of Claim 60, wherein said depth-of-treatment generating means
2 comprises means for generating a depth symbol for each piece of formatted text in the
3 textual document.

1 68. The system of Claim 67, wherein said depth symbol generating means
2 comprises means for selecting a type of each occurrence of a piece of formatted text from a
3 predetermined set of types, means for determining, for each piece of formatted text, the
4 number of each type of occurrence of the formatted text, and means assigning a depth
5 symbol to each piece of formatted text in the textual document based on the total number of
6 occurrences of the formatted text, the number of each type of occurrence of the formatted
7 text, and any verified quotations associated with the piece of formatted text. .

1 69. The system of Claim 68, wherein said depth symbol comprises a symbol
2 having a predetermined number of elements corresponding to the depth-of-treatment value.
3

1 70. The system of Claim 59, wherein the subject matter association means
2 comprises means for identifying sentences adjacent to the occurrences of each piece of
3 formatted text, means for executing a query for each piece of formatted text based on the
4 identified sentences and any verified quotations associated with the piece of formatted text
5 to generate one or more subject matter classifications and a belief score for each potential
6 subject matter classification, and means for selecting one or more subject matter
7 classifications having a belief score above a predetermined threshold to associate with the
8 piece of formatted text.

1 71. The system of Claim 59, wherein said formatted text identification means
2 comprises means for parsing the textual document into one or more paragraphs, means for

3 identifying, for each paragraph, one or more tokens which indicate the presence of a
4 quotation in the sentence, and means for determining, based on the identified tokens,
5 quotations within the textual document. .

1 72. A method for automatically processing a written opinion of a legal case,
2 comprising:
3 identifying each occurrence of a piece of text having a predetermined format, the
4 formatted text being a citation in the written opinion;
5 determining a value for each citation indicating a depth-of-treatment afforded to the
6 particular citation in the textual document;
7 associating a subject matter classification with each citation; and
8 generating a data record containing each citation, the depth-of-treatment value for
9 each citation, and the subject matter classification associated with each citation. .

1 73. The method of Claim 72, wherein each piece of formatted text further
2 comprises a quotation from another prior written decision of a court. .

1 74. The method of Claim 73 further comprises verifying a source for each
2 quotation in the textual document to generate verified quotation information. .

1 75. The method of Claim 74, wherein the verifying further comprises separating a
2 quotation into one or more segment based on the location of a predetermined symbol in the
3 quotation, calculating a word frequency value for each word in each segment, selecting a

4 predetermined number of words in each segment, each selected word having a high word
5 frequency value, generating a template of the quotation indicating the position of each of the
6 selected words in each segment, and comparing the generated template to the text of the
7 quotation in the source textual document. .

1 76. The method of Claim 72 further comprising determining the negative history
2 associated with the citations in the written opinion, the negative treatment determining
3 comprising parsing the textual document to identify occurrences of a predetermined word
4 stem in the textual document indicating that the reasoning of a prior written decision of a
5 court is not proper and verifying said identified word stems to generate a list of overrulings.
6

1 77. The method of Claim 76, wherein said parsing comprises identifying each
2 occurrence of a particular tense of the word stem, gathering words adjacent to each
3 occurrence of the particular tense of the word stem, and selecting a negative treatment phrase
4 based on the gathered words..

1 78. The method of Claim 72 further comprising determining a negative treatment
2 phrase in the textual document comprising parsing the textual document to identify
3 occurrences of a predetermined word in the textual document indicating that the reasoning of
4 a prior written decision of a court is not proper and verifying said identified word stems to
5 generate a list of negative treatments. .

1 79. The method of Claim 73, wherein said citation identification comprises
2 locating each occurrence of a reporter name in the textual document, and gathering words
3 adjacent to the reporter name to identify a beginning and an end of the citation..

1 80. The method of Claim 73, wherein said depth-of-treatment generating
2 comprises generating a depth symbol for each piece of formatted text in the textual
3 document.

1 81. The method of Claim 80, wherein said depth symbol generating comprises
2 selecting a type of each occurrence of a piece of formatted text from a predetermined set of
3 types, determining, for each piece of formatted text, the number of each type of occurrence
4 of the formatted text, and assigning a depth symbol to each piece of formatted text in the
5 textual document based on the total number of occurrences of the formatted text, the number
6 of each type of occurrence of the formatted text, and any verified quotations associated with
7 the piece of formatted text. .

1 82. The method of Claim 81, wherein said depth symbol comprises a symbol
2 having a predetermined number of elements corresponding to the depth-of-treatment value. .
3

1 83. The method of Claim 72, wherein the subject matter association comprises
2 identifying sentences adjacent to the occurrences of each piece of formatted text, executing a
3 query for each piece of formatted text based on the identified sentences and any verified

4 quotations associated with the piece of formatted text to generate one or more subject matter
5 classifications and a belief score for each potential subject matter classification, and
6 selecting one or more subject matter classifications having a belief score above a
7 predetermined threshold to associate with the piece of formatted text. .

1 84. The method of Claim 73, wherein said formatted text identification comprises
2 parsing the textual document into one or more paragraphs, identifying, for each paragraph,
3 one or more tokens which indicate the presence of a quotation in the sentence, and
4 determining, based on the identified tokens, quotations within the textual document. .

1 85. A system for automatically processing a textual document, comprising:
2 means for marking up citations in the textual document, the citations identifying
3 other textual documents relied on in the textual document;
4 means for locating quotations in the textual document which are candidates for being
5 associated with a marked-up citation;
6 means for identifying in the textual document indicating a negative treatment of
7 another textual document;
8 means for generating a depth-of-treatment value for each citation in the textual
9 document, the depth-of-treatment value indicating a potential significance of the piece of
10 text to the reasoning in the textual document;
11 means for determining a source textual document for each candidate quotation in
12 order to verify the origin of the quotation and generate verified quotations;

means for producing a plurality of pieces of text surrounding the citations based on the verified quotations;

means for assigning an abstract to each citation based on the plurality of pieces of text surrounding the citations; and

means for grouping the citations and the verified quotations, the depth-of-treatment values and the abstracts for each citation together into a data record containing information about the reasoning contained in the textual document. .

86. A method for automatically processing a textual document, comprising:

marking up citations in the textual document, the citations identifying other textual documents relied on in the textual document;

locating quotations in the textual document which are candidates for being associated with a marked-up citation;

identifying in the textual document indicating a negative treatment of another textual document;

generating a depth-of-treatment value for each citation in the textual document, the depth-of-treatment value indicating a potential significance of the piece of text to the reasoning in the textual document;

determining a source textual document for each candidate quotation in order to verify the origin of the quotation and generate verified quotations;

producing a plurality of pieces of text surrounding the citations based on the verified quotations;

15 assigning an abstract to each citation based on the plurality of pieces of text
16 surrounding the citations; and
17 grouping the citations and the verified quotations, the depth-of-treatment values and
18 the abstracts for each citation together into a data record containing information about the
19 reasoning contained in the textual document.